

This superficial sketch of Japanese orthography and phonology is an excerpt from:

Vance, Timothy J. 2001. Japanese. *Facts about the World's Languages*, ed. by Jane Garry and Carl Rubino, 345–349. New York & Dublin: H. W. Wilson.

Orthography and Basic Phonology

The Japanese writing system is generally regarded as the modern world's most complex. Roughly speaking, native and Sino-Japanese nouns are written with Chinese characters (*kanji*), as are the invariant portions of most inflected words (verbs and adjectives). Inflectional endings, postpositions, and many adverbs are spelled out in a cursive syllabary (*hiragana*), and non-Chinese loanwords are spelled out in a squareish syllabary (*katakana*). There are, however, many exceptions to these broad generalizations. A single Chinese character often has two or more possible readings, the intended reading being determined by the context. The *hiragana* and *katakana* originated as abbreviated versions of certain Chinese characters that had come to be used as phonograms.

There are two principal systems of Romanization, and a variant of the more popular Hepburn system is used here, with long vowels indicated by double letters (except in place-names mentioned in the text, where macrons are used instead). The chart on the next page shows the individual symbols of *hiragana* and *katakana* in the traditional right-to-left arrangement of top-to-bottom lines. The pronunciation of each symbol is given in Romanization. The many possible Japanese syllables not included in the chart are written by using diacritics and/or by combining two or three symbols.

All consonants listed in the Table I occur initially in CV syllables, and many can also occur initially in CyV syllables. A syllable-final consonant must be one of the two moraic consonants, which count as separate beats in poetic meter. The moraic nasal is always Romanized as *n*, but its pronunciation assimilates to the immediately following sound and varies widely, such as a bilabial [m] before bilabials, as in *onpa* 'soundwave', a nasalized glide [ũ] before *s*, as in *kansa* 'inspection'; before a pause, it is usually uvular [ɴ]. Before a vowel or *y*, a moraic nasal is followed by an apostrophe in Romanization to distinguish it from a syllable-initial alveolar nasal: *an'i* 'easy', *ani* 'older brother'. The moraic obstruent combines with an immediately following obstruent to form a phonetically long obstruent and is Romanized as a copy of that following obstruent, as in *rappa* 'bugle' and *dassen* 'derailment'. Except in recent loanwords, the moraic obstruent occurs only before voiceless obstruents.

The voiceless stops and affricates are unaspirated. Some speakers can substitute a velar nasal [ŋ] for a word-medial voiced velar stop *g* in a large subset of the vocabulary. The alveolar affricates are *ts* [ts] and *z* [dz], but *z* is pronounced as a fricative [z] word-medially unless preceded by the moraic nasal. The alveo-palatal affricates are *ch* [tʃ] and *j* [dʒ]. The fricative *f* [ɸ] was until recently an allophone of *h*, occurring only before *u*, but *f* now appears before all vowels in loanwords, for example, *ferii* 'ferry'. The fricative *h* is pronounced as palatal [ç] before *y* or *i*. The alveo-palatal fricative is *sh* [ɕ]. The liquid is a tap *r* [ɾ], and the velar glide *w* is often described as unrounded [ɰ], although it shows lip compression (not protrusion) in careful pronunciation. Even in recent loanwords, *t* and *d* do not occur before *u*, *s* and *z* do not occur before *y* or *i*, *h* does not occur before *u*, *y* does not occur before *i*, and *w* does not occur before *u*.

The high back vowel *u* is often described as unrounded [ɯ], but like the glide *w* it shows lip compression in careful pronunciation. All five vowels occur both short and long, and a syllable with a long vowel counts as two moras. All two-vowel sequences are possible, although some are quite rare; if the second vowel is *e*, *o*, or *a*, it is in a separate syllable from the first, but if the second vowel is *i* or *u*, it may be the second mora of a long syllable. Short *i* and *u* are frequently devoiced between consonants or between a voiceless consonant and a pause.

Japanese has a pitch-accent system, and the location of accent, manifested as a drop from high to low pitch, is the contrastive feature of the system. The presence or absence of an accent and its location in a word are essential to determining the intonation contour of a phrase containing that word. The possible locations for accent are least constrained in nouns. A noun may be accented on any syllable, or it may be unaccented, as in *ha'shi* 'chopsticks', *hashi* 'bridge' and *hashi* (unaccented) 'edge'. Accent is not marked in Japanese orthography or in Romanization.

Hiragana

ん	わ	ら	や	ま	は	な	た	さ	か	あ
[moraic]n	wa	ra	ya	ma	ha	na	ta	sa	ka	a
		り		み	ひ	に	ち	し	き	い
		ri		mi	hi	ni	chi	shi	ki	i
		る	ゆ	む	ふ	ぬ	つ	す	く	う
		ru	yu	mu	fu	nu	tsu	su	ku	u
		れ		め	へ	ね	て	せ	け	え
		re		me	he	ne	te	se	ke	e
	を	ろ	よ	も	ほ	の	と	そ	こ	お
	(w)o	ro	yo	mo	ho	no	to	so	ko	o

Katakana

ン	ワ	ラ	ヤ	マ	ハ	ナ	タ	サ	カ	ア
[moraic]n	wa	ra	ya	ma	ha	na	ta	sa	ka	a
		リ		ミ	ヒ	ニ	チ	シ	キ	イ
		ri		mi	hi	ni	chi	shi	ki	i
		ル	ユ	ム	フ	ヌ	ツ	ス	ク	ウ
		ru	yu	mu	fu	nu	tsu	su	ku	u
		レ		メ	ヘ	ネ	テ	セ	ケ	エ
		re		me	he	ne	te	se	ke	e
	ヲ	ロ	ヨ	モ	ホ	ノ	ト	ソ	コ	オ
	(w)o	ro	yo	mo	ho	no	to	so	ko	o

Table 1: Consonants

	Bilabial	Alveolar	Alveo-palatal	Velar	Glottal
Stops					
Voiceless	p	t		k	
Voiced	b	d		g	
Fricatives	f	s	sh		h
Affricates					
Voiceless		ts	ch		
Voiced		z	j		
Nasals	m				
Liquid		r			
Glides			y	w	

Table 2: Vowels

	Front	Back
High	i	u
Mid	e	o
Low		a